346609

STIC-ILL

From: Kerr, Kathleen
Sent: Wednesday, May 16, 2001 10:20 AM
To: STIC-ILL
Subject: NPL Orders for 09/311689

Kathleen Kerr, employee # 77468
Art Unit 1652, Rm 10B13
305-1229

I would like to order the following non-patent literature to search on Application number **09/311689**. Thank you very much.

AN 1994:4511 CAPLUS
TI Enhancing the nutritional value of seed crops
AU Beach, Larry R.; Ballo, Barbara
SO **Curr. Top. Plant Physiol. (***1991***), 7(Biosynthesis and Molecular Regulation of Amino Acids in Plants), 229-38**

Kathleen Kerr

Kathleen Kerr, Ph.D.
USPTO Patent Examiner, Art Unit 1652
1911 South Clark Place, Rm 10B13
Arlington, VA 22202
(703) 305-1229

NO 5/16

(7)
•63 & Wa
5/16-R.C

...ivar of rice. Theor Appl Genet **77:** 176–183
...f, Schell J, Spena A (1988) Single genes from *Agrobacterium ...genes* influence plant development. EMBO J **7(9):** 2621–2629
...ul O, Galili G (1992) Increased lysine biosynthesis in tobacco plants that express high levels of bacterial dihydrodipicolinate synthase in their chloroplasts. The Plant Journal **2(2):** 203–209
35. **Shedlarski JG, Gilvarg C** (1970) The pyruvate-aspartic semialdehyde condensing enzyme of *Escherichia coli.* J Biol Chem **245:** 1362–1373
36. **Wallace JC, Galili G, Kawata EE, Cuellar RE, Shotwell MA, Larkins BA** (1988) Aggregation of lysine-containing zeins into protein bodies in Xenopus oocytes. Scienc **240:** 662–664
37. **Widholm JM** (1976) Selection and characterization of cultured carrot and tobacco c lls resistant to lysine, methionine, and proline analogs. Can J Bot **54:** 1523–1529

Biosynthesis and Molecular Regulation of Amino Acids in Plants, *BK Singh, HE Flores, JC Shannon*, eds, Copyright 1992, American Society of Plant Physiologists

Enhancing the Nutritional Value of Seed Crops

Larry R. Beach and Barbara Ballo

(Pioneer Hi-Bred International, Inc.) Department of Biotechnology Research, PO Box 38, Johnston, IA 50131

INTRODUCTION

Swine and poultry growers in developed countries have designed sophisticated feed regimens which result in cost-efficient animal growth. These diets generally combine different plant and synthetic sources of essential nutrients to promote maximal growth at the least cost to the grower. The ultimate source of each component utilized in any geographic area varies and is dependent primarily on regional availability and cost of potential components. Feed formulations must provide for the energy requirements of the animal, supplied as fat and carbohydrate, as well as provide other essential nutrients critical to growth, including a balanced protein source. Grain products such as maize comprise the bulk of carbohydrate contributions to animal feed. Cereals are a relatively inexpensive energy source, but are not the principal protein source in the diet. Additionally, grain proteins are deficient in some of the amino acids (AAs) essential to animal growth: lysine, and to a lesser extent tryptophan and threonine. Dietary protein in animal feed is generally provided by a dicotyledenous seed source such as soy meal. Although dicot seeds are typically deficient in the sulfur-containing AAs, methionine and cysteine, combining dicot seed meal with grain results in complementary AA availability. Neither crop completely compensates for the AA deficiency of the other crop component, however, and feed diets frequently must include supplemental AAs. This need is particularly evident with the exceptional sulfur AA requirement for feather growth by poultry and with the lysine supplementation necessary for maximal swine development.

Methionine supplements to livestock feed, primarily for poultry, amount to 50,000 metric tons per year, at an annual cost to growers of approximately $120M. Twenty thousand metric tons of lysine, valued at $70M, are used to amend deficient swine and poultry diets. It is apparent that a more direct and cost-efficient means of supplying sufficient AAs to these diets would benefit both the grower and the consumer.

Because of its low cost and high carbohydrate content, maize is frequently the major component of animal feeds in the United States. Maize is particularly

deficient in lysine (0.25%). Consequently, the development of high-lysine (0.42%) maize varieties by breeders was heralded as a promising alternative to lysine supplementation of feed diets. Unfortunately, these varieties had low yield and poor kernel characteristics. Although breeders have improved the kernel characteristics, high-lysine maize has a yield penalty which results in less value per acre than with elite maize hybrids. Additionally, these high-lysine maize lines are the result of a recessive mutation, *opaque 2*, which increases the difficulty of making hybrids and requires that they be grown in isolation from normal maize. Other selected maize inbreds which exhibit enhanced methionine content are not likely to be commercially useful either, since the seed from hybrids does not have significantly enhanced methionine levels.

An alternative to relying on inherent variance in AA composition in the crop is to enhance specific AA levels via biotechnology. One method is to express a heterologous protein of favorable AA composition at levels sufficient to obviate feed supplementation. Another is to enhance free AA accumulation via directed control of biosynthetic enzymes, either regulating tissue specificity and degree of expression or modulating enzyme activity by mutation of active and/or regulatory sites of the molecule. We will review here progress in attempts to alter the nutritional value of seed.

EXPRESSION OF HETEROLOGOUS SEED PROTEINS

Transformation and regeneration of fertile dicotyledenous plants has been successful for nearly a decade, while success with monocot transformation is relatively recent. Due to the plethora of data available concerning transgenic expression of dicot seed storage proteins, we will emphasize these in this discussion.

If one wishes to alter the AA composition of the seed, there are several genetic engineering approaches one could attempt. One alternative would be to identify a protein rich in AAs limiting in the target seed. The degree of enrichment necessary to effect a change in AA composition would depend on the amount of expression of the new protein in the transgenic seed, balanced by any potential negative effects expression of a heterologous protein would have on normal protein expression in the developing seed. Finally, when a protein of acceptable AA composition is identified, it must be demonstrated to be acceptable as a dietary supplement, and should be neither antinutritive nor toxic.

Do seed proteins of enriched AA composition exist, and can they be expressed at levels high enough to compensate for deficiencies in host plants? Unfortunately, no seed proteins with a sufficiently high lysine content have been identified yet. A number of seed proteins rich in sulfur AAs have been identified, for example a 10-kD zein of maize (20) is 22.5 mole % methionine and 3.9 mole % cysteine. A 2S protein from Brazil nut is exceptionally rich in sulfur AAs, with 18 mole % methionine and 8 mole % cysteine (2, 11). Another dicot protein related to the 2S class (PA1b) has been identified and cloned from pea and contains 16.2 mole % cysteine (16). Although PA1b has some homology with protease inhibitors, it has not been shown to have protease inhibitor activity. Thus, all three of these proteins would be acceptable for complementary expression in sulfur-AA deficient target plants.

Heterologous expression of seed proteins has been accomplished for the three major dicot seed storage protein classes (11S, 7S, 2S). For a complete description of these protein classes, including the extensive postranslational modifications each exhibits, refer to the review by Higgins (15). It is worth noting that transgenic accumulation of foreign protein has generally been limited to 2% or less of total seed protein. We will review here some of the proteins examined and the potential limiting factors to the extent of expression.

There are a number of factors which may affect heterologous expression. First, dicot seed proteins are encoded by gene families which may contain from a few to nearly 20 members. Individual genes within a family are differentially regulated, some are expressed at different times during seed maturation, and each varies in its contribution to seed protein mass at maturity. Thus, the selection of a specific gene intended either for transformations or for use of its regulatory regions in gene fusions may be critical as it determines the potential extent of expression. Second, postranslational modifications of seed proteins is common. Specifics of modification may not be shared by the heterologous gene and endogenous seed proteins of the target plant, however. It is not clear whether proteolytic cleavage after transport to protein bodies is necessary for stable accumulation. Members of the vicilin-type (7S) classification which are not proteolytically processed in protein bodies in their native environment include phaseolin (*Phaseolus vulgaris*), convicilin (*Pisum sativum*), and conglycinin (*Glycine max*). Each of these proteins exhibits partial though seemingly specific cleavage upon transgenic expression in tobacco or petunia (4, 23, 27). Pea vicilin in tobacco accumulates as component polypeptides identical in mass to those observed in pea seeds as well as several new smaller polypeptides (17). In all examples of transgenic 7S expression supplemental processing was found. This may be attributed to a variation in recognition of specific residues as proteolytic sites, novel proteases in the host plants, or tertiary structure differences in host protein bodies that allow sites to be exposed. Total foreign protein accumulation is not extensive, and although it may result from low copy numbers of transferred genes relative to the copy number in the donor plant, instability resulting from this additional cleavage cannot be ruled out. Pea legumin and sunflower helianthinin, both 11S proteins, have been expressed in tobacco (5, 26). Both have been shown to undergo 11S-type processing, *i.e.*, disulfide linkage within the proproteins followed by cleavage resulting in mature S-linked subunits. Immunologically similar proteins of smaller mass than expected, presumably degradation products, were detected shortly after protein accumulation peaked. Although no quantification of protein expression was presented, it appears from the data supplied that helianthinin comprised a greater proportion of transgenic tobacco seed protein than did pea legumin. Both genes are members of small multigene families, and it may be that the helianthinin

gene used was the most highly expressed of the 11S sunflower genes. Vicilins and legumins rich in essential AA have not yet been identified; however, both classes may provide initial templates for generating modified coding sequences, as will be discussed later.

2S albumins constitute an abundant class of proteins in several dicot plants including canola (*Brassica napus*), Brazil nut (*Bertholletia excelsa*), and *Arabidopsis*. A subset of this class includes proteins exceptionally rich in sulfur-containing AAs, such as Brazil nut protein (BNP). 2S proteins undergo extensive proteolytic processing similar to that of 11S proteins, except mature 2S proteins do not form complex multimeric complexes. Mature protein accumulates as two polypeptide chains of 3 to 5 kD and 8 to 12 kD, cross-linked by oxidation of cysteine residues. 2S proteins from canola (25), *Arabidopsis* (7), and Brazil nut (1, 3, 8, 13) have been expressed transgenically in either tobacco or canola. These proteins were correctly processed and accumulated uniquely in seed protein bodies. Degradation peptides such as those observed in both 7S and 11S transgenic expression were not detected; however, degradation products would likely be too small to be detectable on typical protein gels. Because of its AA composition, the transgenic expression of BNP is of particular importance to this review. The efficacy of chimeric BNP-coding constructs in enhancing methionine content of seeds has been investigated by different laboratories and appears to be at least somewhat dependent on gene construction. Altenbach *et al.* (3) essentially replaced the coding region of a phaseolin gene with the coding sequence of BNP. This translational fusion resulted in BNP accumulation representing up to 8% of transgenic tobacco seed protein and 4% of transgenic canola. Methionine content of these seeds was shown to increase by about 30% in both tobacco (3) and canola (1). A construct utilizing controlling regions derived from an *Arabidopsis* 2S protein gene (*AT2S1*) and including the region encoding the 5' *AT2S1* signal peptide with or without the 5' amino terminal processed fragment was also expressed in tobacco, canola, and *Arabidopsis* (8). Although protein was detectable on immunoblots (less than 1% of salt-soluble protein), expression was not sufficient to allow quantification. A different fusion protein was assembled, consisting of the soybean lectin signal peptide and the propeptide of BNP, with the lectin promoter and the 3' region of *Agrobacterium* gene 7 (13). With this expression cassette BNP was again less than 1% of extracted seed protein. Without knowledge of the relative transcription rate and mRNA or protein stability, it is difficult to conjecture why protein accumulation was low in these experiments. It appears that the *AT2S1* promoter was a poor choice, as other genes in the *Arabidopsis* 2S family are expressed at significantly higher levels (14).

There are a number of approaches to obtaining sufficiently high levels of expression of enriched heterologous proteins such that seed meal will not require supplementation with AAs. Efficient expression cassettes with seed-specific promoters are perhaps the key to good expression. Not only must the gene-controlling regions direct the synthesis of high levels of mRNA, the mRNA must

be translated efficiently into stable protein. The phaseolin-promoted Brazil nut 2S expression cassette is an example of an effective chimeric seed-specific gene.

There are other variables which could be important for high levels of protein accumulation, but there are insufficient data to evaluate their significance. Will the insertion of multiple copies of an expression cassette result in higher levels of expression? Are the levels of free AAs limiting? Is there an upper limit to the accumulation of foreign proteins in the seed because of an inherent physiological limit of seed protein content? This physiological limitation would limit the levels of foreign proteins because they replace native seed proteins and the viability of the seed may be compromised with the concomitant native seed protein loss.

Currently, the concept of expressing high levels of methionine-rich seed-storage proteins is a viable one. Chimeric expression cassettes can be incorporated into plants which express easily detectable levels of the seed-storage protein. These genes are stably transmitted to progeny. The accumulated proteins do not appear to affect seed viability. Tobacco and canola seed containing the Brazil nut 2S protein have an enhanced methionine content. The next challenge is to obtain similar results in an important feed crop deficient in methionine, such as soybean.

MODIFICATION OF EXISTING PROTEINS

One approach to altering the AA content of seeds is to alter the sequence of existing seed protein genes and reintroduce the altered gene into plants. These alterations could be either codon changes to specify different AAs or the insertion of a new coding sequence to obtain an increased molar percentage of the desired AA. Since storage proteins generally undergo postranslational processing, are targeted to storage vacuoles, and are assembled into multimeric complexes, the secondary and tertiary structure of the storage protein is critical. Any alteration of the primary sequence could easily result in a disruption of the above processes. Thus, any change in the AA sequence is best made with the knowledge of the higher order folding of the protein and the significance of targeting and processing sites. Recently the first X-ray structural analysis of a seed storage protein was accomplished (21). With complete three-dimensional structural information one can begin to predict where changes and additions can be made without disturbing the structure. The alterations in storage proteins summarized here were accomplished without the benefit of knowing the likely three-dimensional structure.

The first report of the expression of a modified storage protein in seeds of a transgenic plant was by Hoffman *et al.* (19). They inserted the coding sequence for a 15-AA peptide from a maize 15-kD zein into the β-phaseolin storage protein from *P. vulgaris*. The resulting phaseolin (himet) contained three times the number of methionine residues as unmodified phaseolin. Although transgenic plants containing the unmodified and himet phaseolin genes had similar levels of the respective mRNAs, unmodified phaseolin accumulated to levels 500 times those of himet phaseolin. Although some of the himet phaseolin appeared to be processed and

assembled into trimers normally, the majority was postranslationally unstable. Unlike unmodified phaseolin, himet phaseolin did not accumulate in protein bodies in transgenic tobacco seeds. Although Hoffman *et al.* (19) had attempted to insert a peptide with a predicted secondary structure which would match the structure surrounding the insertion site, the insertion site was an unlucky choice based on the structure predicted by Lawrence *et al.* (21). Their X-ray diffraction analysis of the three-dimensional structure of phaseolin identifies the insertion site as "part of a major structural element of the phaseolin trimer" (21).

Since it had been shown that the region of the *Arabidopsis* 2S albumin between the sixth and seventh cysteine residues tolerated modifications (29), DeClercq *et al.* (8) replaced a 23-AA coding segment in this region with three different high-methionine coding fragments. The resulting chimeric proteins contained insertions of peptides 7, 18, and 24 AA long, which increased the number of methionine residues in the protein by four, seven, and 11, respectively. These modified *Arabidopsis* 2S genes were transformed into *Arabidopsis thaliana*, *B. napus*, and *Nicotiana tabacum*. The modified proteins were expressed in seeds of each of the above species; however, only an indirect measurement of the modified protein accumulation in transgenic *Arabidopsis* seeds was presented. In the better expressing transgenic plants each of the modified proteins accounted for up to 1 to 2% of the high-salt extracted seed proteins (8). The authors predicted that the above expression level for the most highly modified 2S albumin would result in only a 5% increase in methionine content of *Arabidopsis* seed, and suggested that one explanation for the low level of 2S protein accumulation was the choice of promoter (14). Although a poor promoter would contribute to poor accumulation, it is not possible to rule out postranslational instability of the protein. While this approach did not result in a significant increase in methionine in the seed, it did illustrate the capacity to modify and express 2S protein in the seed.

Ohtani *et al.* (24) altered four different codons in the gene which codes for a 19-kD α-zein. Each alteration converted a nonlysine codon to a lysine codon, thereby producing four different modified α-zein sequences. Each modified zein coding region was inserted into an expression cassette with a β-phaseolin promoter and a β-zein polyadenylation signal and subsequently transformed into tobacco. The level of α-zein mRNA in the transgenic seed varied from 0.05 to 2.5% of the polyA containing RNA; there was no detectable difference in mRNA levels between modified and unmodified α-zein. The level of accumulation of the unmodified α-zein was up to 0.003% of the total seed protein, and the modified zeins were detected at levels 30- to 300-fold lower in transgenic seed. In these experiments the accumulation of modified or unmodified α-zein was much less than expected based on the amount of mRNA present in developing seeds. *In vivo* pulse-labelling experiments indicated that the newly synthesized protein had a very short half-life. Thus, it appears that α-zein, unlike β-zein (18), is unstable in transgenic tobacco seeds (24).

One interpretation of the experiments resulting in low levels of altered phaseolin or zein is that it will be impossible to alter storage proteins and have

significant levels of expression in transgenic seed. We would not agree with this interpretation. It is highly probable that these first generation storage protein alterations were largely unsuccessful because the alterations were made in regions critical to targeting, processing, or assembly. With further understanding of the tertiary structure and sequences critical to assembly and accumulation such as has been attained with β-conglycinin (22), it will be possible to engineer proteins predictively.

MODIFICATION OF AMINO ACID BIOSYNTHESIS

Heretofore in this discussion the enhancement of specific AA levels has been addressed by altering the levels of proteins containing the desired AA. An alternative approach is to increase the levels of free AA. Since considerable attention is given to this area by other contributors to this volume, we will summarize only recent work on the "aspartate family" of AAs. Methionine, threonine, isoleucine, and lysine are aspartate-derived; the first step in their synthesis is dependent on the phosphorylation of aspartate by aspartate kinase (AK). Two steps beyond AK is a branch-point in the pathway which is the first committed step to lysine synthesis, dihydrodipicolinic acid synthase (DHDPS). As with many metabolic pathways, regulation of the enzymatic synthesis of product is controlled allosterically by endproduct feedback inhibition. Aspartate kinase is feedback-inhibited by threonine and lysine, while DHDPS is feedback-inhibited only by lysine (6). Thus, the I_{50} of DHDPS determines the level of lysine accumulation; in plants the I_{50} is between 10 and 50 μM lysine (6). Many attempts have been made to obtain mutants which are no longer feedback-inhibited by lysine. These mutants were selected by mutagenizing and then selecting for resistance to high concentrations of lysine and threonine or lysine analogs such as S-2-aminoethyl cysteine. Plants were obtained containing elevated levels of free threonine and lysine; they possess lysine-insensitive AK and DHDPS, respectively (9, 10, 12). For the ultimate purpose of obtaining seeds with higher levels of essential AA, mutants obtained in this fashion are unsatisfactory, since there appears to be little to no change in the AA content of the seeds.

Because bacterial DHDPS is not feedback-inhibited as strongly by lysine (I_{50} = 1 mM), transgenic plants expressing the bacterial DHDPS have higher levels of free lysine (12). The *Escherichia coli dap*A gene coding for DHDPS was transformed into *N. tabacum* SR1 with a construct containing a cauliflower 35S promoter, a chloroplast transit peptide from a ribulose bisphosphate carboxylase small subunit gene from pea, the *dap*A coding region, and the *nos* polyadenylation sequence. The levels of lysine in the leaves of these plants were up to 200-fold those in normal tobacco leaves (12). Similar results were obtained by Shaul and Galili (28); *N. tabacum* transformed with the *E. coli dap*A gene had 40-fold elevated levels of lysine. They showed that elevated lysine levels were obtained when the bacterial DHDPS was targeted to the chloroplast, where plant DHDPS normally functions. Accumulation of DHDPS in the cytoplasm did not result in any increase in lysine (28). Seed lysine content was not given in these reports (12, 28).

CONCLUSION

The goal of obtaining increased levels of essential AA in plants has been approached by expressing storage proteins containing high levels of the desired AA or by modifying the AA biosynthetic enzymes. Although significant increases in the free lysine levels in leaves have been obtained by selection for DHDPS mutants or by expressing the *E. coli* DHDPS in plants, it remains to be shown that these alterations can increase seed lysine content. Experiments to increase the levels of specific AA in storage proteins by altering known proteins have met with only marginal success. It is likely that altering a specific protein sequence will be more successful with increased knowledge of the peptide sequences critical to processing, intracellular targeting, and assembly into stable forms.

The most successful approach to altering seed AA content has been the expression of the Brazil nut 2S storage protein containing high levels of methionine and cysteine (1, 3). Because this protein has a sulfur AA content of 26 mole % and it was expressed at high levels in the seed, the impact on the seed AA content was significant. Since the Brazil nut 2S protein does not always accumulate to such high levels in transgenic seed (14), it is clear that choosing the right storage protein is not the only critical factor. The key to this success is likely a combination of the compatibility of the storage protein with the target seeds as well as the use of efficient expression cassettes (1, 3).

LITERATURE CITED

1. **Alt nbach SB, Kuo CC, Staraci LC, Pearson KW, Wainwright C, Georgescu A, T wns nd J** (1992) Accumulation of a Brazil nut albumin in seeds of transgenic canola results in enhanced levels of seed protein methionine. Plant Mol Biol **18:** 235–245
2. **Altenbach SB, Pearson KW, Leung FW, Sun SSM** (1987) Cloning and sequence analysis of a cDNA encoding a Brazil nut protein exceptionally rich in methionine. Plant Mol Biol **8:** 239–250
3. **Alt nbach SB, Pearson KW, Meeker G, Staraci LC, Sun SSM** (1989) Enhancement of the methionine content of seed proteins by the expression of a chimeric gene encoding a methionine-rich protein in transgenic plants. Plant Mol Biol **13:** 513–522
4. **Beachy RN, Chen Z-L, Horsch RB, Rogers SG, Hoffmann NJ, Fraley RT** (1985) Accumulation and assembly of soybean β-conglycinin in seeds of tranformed petunia plants. EMBO J **4:** 3047–3053
5. **B gu MA, V nder Haar RA, Nuccle ML, Griffing LR, Thomas TL** (1990) Developmentally regulated expression of a sunflower 11S seed protein gene in transgenic tobacco. Mol Gen Genet **222:** 49–57
6. **Bryan JK** (1980) Synthesis of the aspartate family and branched-chain AAs. *In* BJ Miflin, ed, The Biochemistry of Plants, Vol 5. Academic Press, New York, pp 403–452
7. **DeClercq A, Vandewiele M, De Rycke R, Van Damme J, Van Montagu M, Krebb rs E, Vandekerckjove J** (1990) Expression and processing of an *Arabidopsis* 2S albumin in transgenic tobacco. Plant Physiol **92:** 899–907
8. **DeClercq A, Vandewiel M, Van Damme J, Guerche P, Van Montagu M, Vandekerckhove J, Kr bbers E** (1990) Stable accumulation of modified 2S albumin seed storage proteins with higher methionine contents in transgenic plants. Plant Physiol **94:** 970–979
9. **Frankard V, Ghislain M, Negrutiu I, Jacobs M** (1991) High threonine producer mutant of *Nicotiana sylvestris* (Spegg. and Comes). Theor Appl Genet **82:** 273–282
10. **Frisch DA, Tommey AM, Gengenbach BG, Som rs DA** (1991) Direct genetic selection of a maize cDNA for dihydrodipicolinate synthase in an *Escherichia coli dapA*⁻ auxotroph. Mol Gen Genet **228:** 287–293
11. **Gander ES, Holmstroem K-O, De Paiva GR, De Castro LAB, Carneiro M, Grossi de Sa F** (1991) Isolation, characterization and expression of a gene coding for a 2S albumin from *Bertholletia excelsa* (Brazil nut). Plant Mol Biol **16:** 437–448
12. **Glassman KF, Barnes LJ, Pilacinski WP** (1989) Method of inducing lysine overproduction in plants. International Patent Application No PCT/US89/01309
13. **Guerche P, De Almeida ERP, Schwarztein MA, Gander E, Krebbers E, Pelletier G** (1990) Expression of the 2S albumin from *Bertholletia excelsa* in *Brassica napus*. Mol Gen Genet **221:** 306–314
14. **Guerche P, Tire C, Grossi de Sa F, DeClercq A, Van M ntagu M, Krebbers E** (1990) Differential expression of the *Arabidopsis* 2S albumin genes and the effect of increasing gene family size. Plant Cell **2:** 469–478
15. **Higgins TJV** (1984) Synthesis and regulation of major proteins in seeds. Ann Rev Plant Physiol **35:** 191–221
16. **Higgins TJV, Chandler PM, Randall PJ, Spencer D, B ach LR, Blagrov RJ, Kortt AA, Inglis AS** (1986) Gene structure, protein structure, and regulation of the synthesis of a sulfur-rich protein in pea seeds. J Biol Chem **261:** 11124–11130
17. **Higgins TJV, Newbigin EJ, Spencer D, Llewellyn DJ, Craig S** (1988) The sequence of a pea vicilin gene and its expression in transgenic tobacco plants. Plant Mol Biol **11:** 683–695
18. **Hoffman LM, Donaldson DD, Bookland R, Rashka K, H rman EM** (1987) Synthesis and protein body deposition of maize 15-kD zein in transgenic tobacco seeds. EMBO J **6:** 3213–3221
19. **Hoffman LM, Donaldson DD, Herman EM** (1988) A modified storage protein is synthesized, processed, and degraded in the seeds of transgenic plants. Plant Mol Biol **11:** 717–729
20. **Kirihara JA, Petri JB, Messing J** (1988) Isolation and sequence of a gene encoding a methionine-rich 10-kD zein protein from maize. Gene **71:** 359-370
21. **Lawrence MC, Suzuki E, Varghese JN, Davis PC, Van Donkelaar A, Tulloch PA, Colman PM** (1990) The three-dimensional structure of the seed storage protein phaseolin at 3Å resolution. EMBO J **9:** 9–15
22. **Lellevre J-M, Dickinson CD, Dickinson LA, Nielsen NC** (1992) Synthesis and assembly of soybean β-conglycinin *in vitro*. Plant Mol Biol **18:** 259–274
23. **Newbigin EJ, de Lumen BO, Chandler PM, Gould A, Blagrove RJ, March JF, Kortt AA, Higgins TJV** (1990) Pea convicilin: structure and primary

sequence of the protein and expression of a gene in the seeds of transgenic tobacco. Planta **180:** 461–470

24. **Ohtani T, Galili G, Wallace JC, Thompson GA, Larkins BA** (1991) Normal and lysine-containing zeins are unstable in transgenic tobacco seeds. Plant Mol Biol **16:** 117–128
25. **Radke SE, Andrews BM, Moloney MM, Crouch ML, Kridl JC, Knauf VC** (1988) Transformation of *Brassica napus* L. using *Agrobacterium tumefaciens*: developmentally regulated expression of a reintroduced napin gene. Theor Appl Genet **75:** 685–694
26. **Rerie WG, Whitecross M, Higgins TJV** (1991) Developmental and environmental regulation of pea legumin genes in transgenic tobacco. Mol Gen Genet **225:** 148–157
27. **Sengupta-Gopalan C, Reichert NA, Barker RF, Hall TC, Kemp JD** (1985) Developmentally-regulated expression of the bean β-phaseolin gene in tobacco seed. Proc Natl Acad Sci USA **82:** 3320–3324
28. **Shaul O, Galili G** (1992) Increased lysine synthesis in tobacco plants that express high levels of bacterial dihydrodipicolinate synthase in their chloroplasts. Plant J **2:** 203–209
29. **Vandekerckhove J, Van Damme J, Van Lijsebettens J, Botterman J, De Block M, Vandewiele M, DeClercq A, Leemans J, Van Montagu M, Krebbers E** (1989) Enkephalins produced in transgenic plants using modified 2S seed storage proteins. Bio/Technology **7:** 929–932

Biosynthesis and Molecular Regulation of Amino Acids in Plants, B.K. Singh, Shannon, eds, Copyright 1992, American Society of Plant Physiologists

Secondary Metabolite Profiles of Crucifer Seeds: Biogenesis, Role, and Prospects for Directed Modification

Clint C. Chapple and Brian E. Ellis

DOE Plant Research Laboratory, Michigan State University, East Lansing, MI 48824-1312; and Department of Plant Science, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4

INTRODUCTION

The family Cruciferae has had a long history of importance to human society, as a source of food, condiments, and medicinal preparations. It is the oil-producing ability and chemistry of the crucifers that has been responsible for this prominence, especially their ability to synthesize the mustard oil glucosides (now referred to as glucosinolates) that typify members of this and a few related families. In addition to glucosinolates, however, two other specialized metabolites have had a more recent impact on the utilization of these species: erucic acid and sinapine (3,5-dimethoxy-4-hydroxycinnamoyl choline). The ability of the oilseed *Brassica* species, notably *B. napus* and *B. campestris*, to produce high levels of oil when cultivated in temperate climatic regions has made them attractive crops for North American and European producers. *Brassica* storage lipids, however, normally contain a high level of the toxic fatty acid, erucic acid. Before an acceptable edible oil could be developed, it was necessary for rapeseed breeders to eliminate erucic acid from the fatty acid profiles of the seed oil and to obtain major reductions in the levels of seed glucosinolates. Both goals were achieved, but while the erucic acid problem has been satisfactorily resolved, the levels of both glucosinolates and sinapine in commercial cultivars require further improvement.

GLUCOSINOLATES

Glucosinolates as a class can be represented as sulfated thiohydroximates that have been stabilized by glucosylation of the thiol group (Fig. 1A). They are biosynthetically derived from a limited number of amino acids through pathways that have been established largely by radiotracer studies (25). The flavor components that make the crucifers so distinctive are the products of rapid structural rearrangements that occur when the aglycones of the glucosinolates are released